

AMENDMENT TO THE CLAIMS

1. (currently amended) A method of processing a body of text to generate compression options, comprising:

performing a ~~linguistic~~-syntactic analysis on the body of text to obtain a linguistic output indicative of linguistic components of the body of text wherein the syntactic analysis includes a tree having non-terminal nodes representing multi-word portions of the body of text and terminal nodes indicative of words in the body of text;

after performing the linguistic analysis, automatically generating a plurality of compression options for each of a plurality of different portions of the body of text to compress the body of text based on the linguistic output, each of the compression options comprising a different compressed form of an instance of the portion in the body of text wherein automatically generating comprises: and

automatically subjecting the portions of the body of text to different sets of compression rules to obtain the plurality of compression options, by subjecting each portion of the body of text to the different sets of compression rules in a predetermined order such that the compression options reflect varying degrees of compression of a same portion of the body of text, wherein both the non-terminal nodes and terminal nodes are subjected to the compression rules;

generating a compression identifier attribute indicative of at least one of the sets of compression rules to which the portion of the body

of text is subjected; and
generating a ShortForm attribute indicative of a
compressed form of the portion of the body of text
by applying the set of compression rules to the
syntactic analysis; and

selecting one of the plurality of compression options for each of the plurality of different portions of the body of text to output a compressed form of the body of text.

2. Canceled.
3. Canceled.
4. Canceled.
5. Canceled.

6. (Currently amended) The method of claim 5-1 wherein generating automatically a plurality of compression options comprises:

generating a case normalized attribute, based on the ShortForm attribute, indicative of a CaseNormalizedForm of the ShortForm attribute.

7. (previously amended) The method of claim 6 wherein generating automatically a plurality of compression options comprises:

generating a compression attribute indicative of a further compressed form of the case normalized attribute.

8. (original) The method of claim 7 wherein generating a compression attribute comprises:

applying letter removal rules to the case normalized attribute to remove letters based on a predetermined location of the letters in the CaseNormalizedForm.

9. (previously amended) The method of claim 8 wherein generating automatically a plurality of compression options comprises:

generating a LongForm attribute that reflects substantially no compression of the portion of the body of text.

10. (previously amended) The method of claim 9 wherein one ShortForm attribute comprises a word substitution based on a dictionary look-up and wherein generating automatically a plurality of compression options comprises:

setting the case normalized attribute and the compression attribute to the ShortForm attribute.

11. Canceled.

12. (Currently Amended) The method of claim ~~11~~1 wherein the linguistic analysis further comprises, prior to performing the syntactic analysis:

performing a lexical analysis on the body of text; and
performing a morphological analysis on the body of text.

13. (Currently Amended) The method of claim ~~5~~1 wherein generating the ShortForm attribute comprises:

normalizing dates to a numerical form.

14. (Currently Amended) The method of claim ~~5~~1 wherein generating the ShortForm attribute comprises:

normalizing offset dates to a numerical form, based on a date that the body of text was authored.

15. (Currently Amended) The method of claim ~~5~~1 wherein generating the ShortForm attribute comprises:

maintaining symbol-sensitive text fragments in uncompressed

form.

16. (original) The method of claim 15 wherein maintaining symbol-sensitive text fragments comprises:

maintaining text fragments that, cannot be accurately understood unless maintained fully in-tact, in uncompressed form.

17. (original) The method of claim 16 wherein maintaining text fragments comprises:

maintaining uniform resource locators and electronic mail addresses in uncompressed form.

18. Canceled.

19. Canceled.

20. Canceled.

21. Canceled.

22. Canceled.

23. Canceled.

24. Canceled.

25. Canceled.

26. Canceled.

27. Canceled.

28. Canceled.

29. Canceled.

30. Canceled.

31. Canceled.